УДК 681.324:519.21

Е.И. БОБЫР, д-р техн. наук, ХВУ (г. Харьков),
Л.Ч.УГОРЕНКО, ХВУ (г. Харьков)

МЕТОД КОНТРОЛЯ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ВЫЧИСЛИТЕЛЬНОГО ПРОЦЕССА В МНОГОЗАДАЧНЫХ СРЕДАХ

У статті розглядається проблема організації ефективної роботи програмної системи з жорсткими часовими обмеженнями. Запропонований метод контролю за ходом обчислень заснований на виявленні тренда середнього значення випадкового часу (в умовах багатозадачності) періодичного виконання певних програмних модулів. Приведені рекурентні вирази дозволяють одержувати в реальному масштабі часу прогнозні оцінки, на основі яких здійснюється адаптивне управління обчислювальним процесом.

In article the problem of the organization of effective time-restricted program system work is considered. The offered method of calculations monitoring is based on revealing of a trend of average value of stochastic time performance (in multitask environment) periodically called program modules. The presented recurrent expressions allow in real time to receive forecast estimations, on the basis of which adaptive management of computing process is carried out.

Постановка проблемы. Отличительной особенностью надежных программных систем является организация устойчивого вычислительного процесса, адаптивного к различного рода возмущениям.

Под адаптацией понимается [1] процесс изменения структуры, алгоритмов и параметров системы на основе информации, получаемой в процессе ее функционирования для достижения оптимального по принятому критерию состояния системы при наличии неопределенности и изменчивости условий работы во взаимодействии с внешней средой.

В основе адаптации лежит накопление данных о поведении системы и внешних условиях и их использование для улучшения избранного показателя качества. Процесс накопления и обработки информации связан с затратами времени, что в итоге приводит к некоторому запаздыванию при выработке решений, компенсирующих возмущающие воздействия. Это может существенно снижать эффективность работы программных систем в реальном масштабе времени. Поэтому для построения адаптивных программных систем задача прогнозирования состояний или поведения системы и внешней среды является актуальной.

Решение указанной задачи для программных систем, функционирующих в многозадачных средах, осложняется [2]:

– случайным динамическим изменением количества обрабатываемых системой разнотипных заявок, определяемых, как правило, внешней средой и активирующих различные процессы;

– наличием в каждый конкретный момент времени разного количества и объема вычислительных ресурсов, доступных для использования активным процессом;

– неопределенностью требуемых при обработке даже конкретной заявки ресурсных затрат (например, при решении итерационной задачи для достижения необходимой точности результатов, а также наличии ветвлений, циклов, прерываний, конфликтов и т.д.).

В таких условиях своевременное выявление критической загрузки системы (в случае возрастания до предельного числа обрабатываемых заявок; при появлении или активации в многозадачной системе дополнительных процессов, потребляющих ресурсы, и т.п.) для адаптивных программных систем реального времени должно осуществляться путем использования эффективных методов контроля за параметрами вычислительного процесса, и в первую очередь за его временными характеристиками.

Анализ литературы. Управление выполнением параллельных вычислительных процессов, основанное на динамическом уточнении статических прогнозов времени выполнения программных модулей, предлагается в [2, 3]. При этом время исполнения каждого модуля рассматривается как случайная величина, распределенная по экспоненциальному закону, а в основе решения на адаптацию лежит простое сравнение измеренных реальных оценок с рассчитанным ранее критическим порогом.

Такое допущение, как показано в [4, 5], является весьма сомнительным. Кроме того, предложенный метод не учитывает характерную особенность многих программных систем – периодичность выполнения отдельных задач – и связанную с ней возможность анализа значений нескольких измерений за некоторый период наблюдений с целью выявления динамики изменения временных параметров исполняемых программ.

В [6] рассмотрены строгая постановка задачи обнаружения изменения свойств временных рядов, критерии и методы ее решения:

1. Методы последовательного (в темпе поступления информации) анализа:

– алгоритм кумулятивных сумм;

– алгоритм на основе критерия Неймана-Пирсона;

– алгоритм экспоненциального сглаживания.

2. Апостериорные методы:

– алгоритм на основе критерия Байеса;

– алгоритм модели авторегрессии – скользящего среднего.

Эффективность применения данных методов определяется наличием априорной информации о виде и параметрах распределения наблюдаемой случайной величины, характером ее изменения и значением устанавливаемого критического значения для выбранного параметра. Кроме того, рассмотренные методы определяют «разладку» процесса (превышение заданного порога) постфактум и обладают определенным временем запаздывания ее обнаружения.

Цель статьи. В статье предлагается новый метод контроля за ходом вычислительного процесса, который лишен рассмотренных недостатков и позволяет в реальном масштабе времени осуществить прогноз состояния программной системы на основе выявления тренда (т.е. изменения во времени) среднего значения времени выполнения некоторого периодически вызываемого программного модуля.

Основной раздел. Оценки времени исполнения программы, получаемые, например, с помощью измерительных мониторов [5, 7], представим в виде ряда

$$y(t_i) = f(t_i) + n(t_i), \; i = 1, 2, \ldots, N,$$

где N – число проведенных измерений; $f(t_i)$ – детерминированная функция, определяющая характер изменения временных оценок (тренд); $n(t_i)$ – стационарная случайная функция, значение которой обусловлено стохастической природой времени выполнения программы, внешними воздействиями, ошибками измерения и т.д.

Дальнейшие выкладки основаны на допущении, что моменты измерений образуют пуассоновский поток событий с интенсивностью λ, а случайные величины $n(t_i)$ независимы и одинаково распределены с математическим ожиданием, равным нулю.

Аппроксимируем $f(t_i)$ линейной функцией

$$f(t_i) = a + b\frac{\lambda t_i}{N}, \tag{1}$$

при этом параметры a и b имеют одинаковую размерность, совпадающую с размерностью процесса $y(t_i)$. Тогда задача выделения тренда временного ряда

$$y_i = y(t_i) = a + b\frac{\lambda t_i}{N} + n(t_i)$$

сводится к задаче получения неизвестных коэффициентов a и b, рассмотренной в [8].

Поскольку $t_i = \tau_1 + \tau_2 + \ldots + \tau_i$, где τ – интервал между проводимыми замерами, и для принятой модели измерений $P(\tau) = \lambda e^{-\lambda\tau}$,

$$M(t_i) = iM(\tau) = \frac{i}{\lambda},$$

$$M(y_i) = a + b\frac{\lambda}{N}M(t_i) = a + b\frac{i}{N}. \qquad (2)$$

Оценки $\hat{a}$ и $\hat{b}$ могут быть найдены методом наименьших квадратов по критерию:

$$\sum_{i=1}^{N}\left(y_i - \hat{a} - \hat{b}\frac{i}{N}\right)^2 \to \min_{a,b}.$$

Вычисляя производные по $\hat{a}$ и $\hat{b}$ и приравнивая их нулю, получаем следующую систему уравнений:

$$\begin{cases} \sum\limits_{i=1}^{N}\left(y_i - \hat{a} - \hat{b}\frac{i}{N}\right) = 0, \\ \sum\limits_{i=1}^{N}\frac{i}{N}\left(y_i - \hat{a} - \hat{b}\frac{i}{N}\right) = 0, \end{cases}$$

решение которой позволяет получить выражения

$$\hat{a} = \frac{2(2N+1)}{N(N-1)}L_0 - \frac{6}{N(N-1)}L_1,$$

$$\hat{b} = -\frac{6}{N-1}L_0 + \frac{12}{N^2-1}L_1, \qquad (3)$$

где

$$L_0 = \sum_{i=1}^{N} y_i, \quad L_1 = \sum_{i=1}^{N} iy_i.$$

Оценка $\hat{b}$, как видно из (1), определяет наклон прямой, аппроксимирующей имеющиеся значения, и позволяет сделать вывод об увеличении (уменьшении) среднего значения времени выполнения программы за N измерений.

Для определения доверительного интервала для оценки $\hat{b}$ необходимо знание ее закона плотности распределения и значения дисперсии $D[n(t_i)]$. В [8] показано, что в качестве несмещенной оценки дисперсии может быть взята статистика

$$s^2 = \gamma_0 L_0^2 + \gamma_1 L_0 L_1 + \gamma_2 L_1^2 + \gamma_3 L_2,$$

где

$$L_2 = \sum_{i=1}^{N} y_i^2 ,$$

а коэффициенты, обеспечивающие несмещенность, определяются выражениями:

$$\gamma_0 = -\frac{2(10N^2 + 13N + 9)}{N(N-1)(N-2)(5N+1)} ,$$

$$\gamma_1 = \frac{60(N+1)}{N(N-1)(N-2)(5N+1)} ,$$

$$\gamma_2 = -\frac{60}{N(N-1)(N-2)(5N+1)} ,$$

$$\gamma_3 = -\frac{(5N^2 - 4N + 3)}{(N-1)(N-2)(5N+1)} .$$

При больших значениях N величина $\frac{s^2(N-2)}{\sigma^2}$ имеет распределение χ^2 с N–2 степенями свободы, а закон распределения оценки $\hat{b}$ асимптотически приближается к нормальному, что позволяет определить значения доверительного интервала по формуле

$$b_{1,2} = \frac{1}{1 - \frac{6}{5}\frac{g_\alpha^2}{N}} \left(\hat{b} \mp \sqrt{\frac{6}{5}\frac{g_\alpha^2}{N}\hat{b}^2 + 12\frac{s^2(N-2)}{\chi^2_{N-2,\alpha}}\frac{g_\alpha^2}{N}\left(1 - \frac{6}{5}\frac{g_\alpha^2}{N}\right)} \right), \qquad (4)$$

где 2α – заданный уровень значимости; g_α – соответствующий квантиль нормального распределения; $\chi^2_{N-2,\alpha}$ – соответствующий квантиль распределения χ^2.

Таким образом, при практической реализации предложенного метода обнаружения тренда необходима организация «скользящего окна» измеряемых значений $y_1, \ldots, y_N$ времени выполнения периодически вызываемой программы и рекуррентный расчет на k-м шаге используемых в (3), (4) базовых значений

$$L_0^{(k)} = L_0^{(k-1)} - y_1 + y_k ,$$
$$L_1^{(k)} = L_1^{(k-1)} - L_0^{(k-1)} + y_k N , \qquad (5)$$
$$L_2^{(k)} = L_2^{(k-1)} - y_1^2 + y_k^2 .$$

Прогнозное значение времени выполнения программы на $k+i$ момент времени может быть получено согласно выражению (2).

Следует учитывать, что достоверность и точность решения задачи прогнозирования существенно зависят от количества реализаций N, которые затрачены на получение статистического прогноза. Таким образом, возникает проблема поиска компромисса между необходимостью увеличения затрат времени и памяти на получение и хранение необходимого числа обрабатываемых данных, и необходимостью уменьшения потребления ресурсов ЭВМ, исходя из требований обработки информации в реальном масштабе времени.

Для решения данной проблемы предлагается использование следующих рекуррентных соотношений, зависящих только от нового поступившего измерения y_k:

$$L_0^{(k)} = L_0^{(k-1)} - (\beta L_0^{(k-1)} / N + (1-\beta)(\hat{a}-\hat{b})) + y_k,$$
$$L_1^{(k)} = L_1^{(k-1)} - L_0^{(k-1)} + y_k N, \quad (6)$$
$$L_2^{(k)} = L_2^{(k-1)} - (\beta L_2^{(k-1)} / N + (1-\beta)(\hat{a}-\hat{b})) + y_k^2.$$

В предлагаемых выражениях вычитание значений y_1 и y_1^2, относящихся к началу «скользящего окна», заменено вычитанием величин, определяемых средним значением статистик L и обратным прогнозным значением y_{k-N}, вычисляемым по (2) на основе текущих параметров $\hat{a}$ и $\hat{b}$. Коэффициент β, как в алгоритмах экспоненциального сглаживания, определяет «степень доверия» к прогнозному и среднему значению.

Пример измерений времени выполнения реальной программы, оценка его среднего значения и рассчитанный доверительный интервал для оценок $\hat{b}$ наклона прямой с использованием выражений (5) приведены на рис. 1 – 3, а с применением формул (6) – на рис. 4.

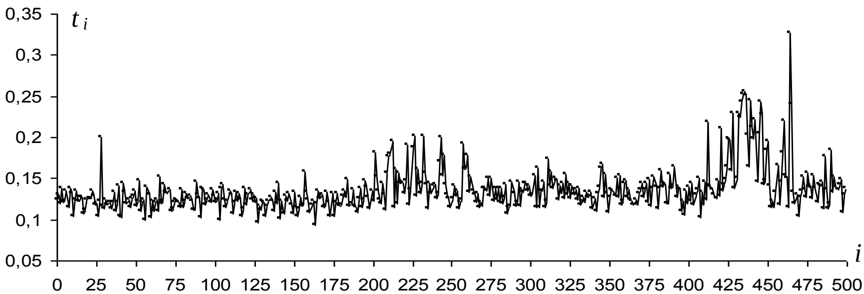

Рис. 1. Замеры времени выполнения (сек) периодически вызываемой программы в многозадачной среде Windows (500 измерений)

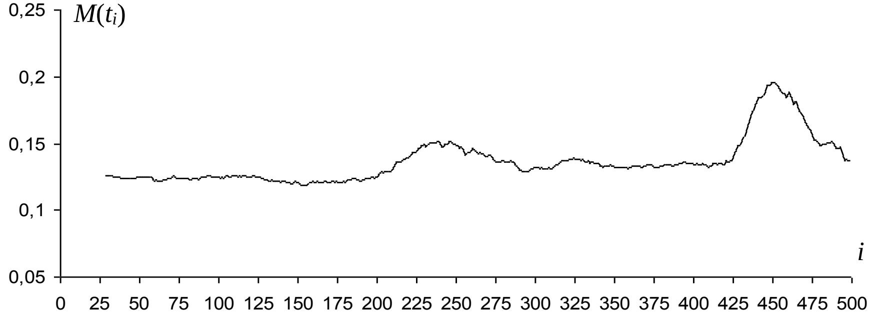

Рис. 2. Изменение среднего времени исполнения программы («скользящее окно» в N=30 измеренных значений)

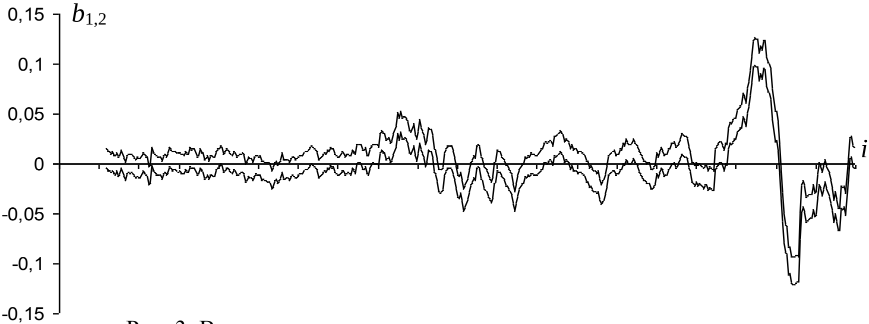

Рис. 3. Верхняя и нижняя границы доверительного интервала для оценок $\hat{b}$ («скользящее окно», уровень значимости $\alpha = 0{,}05$)

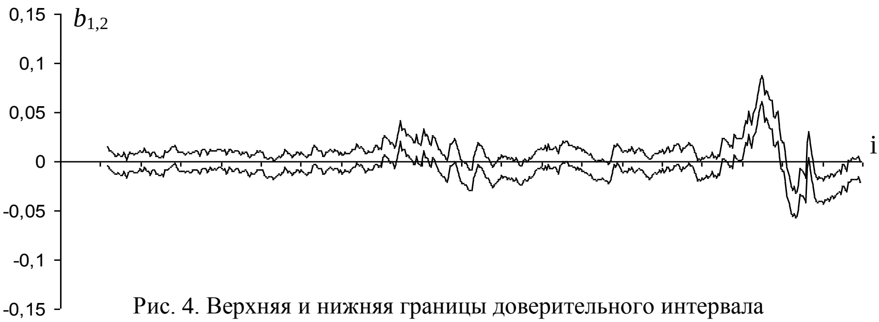

Рис. 4. Верхняя и нижняя границы доверительного интервала для оценок $\hat{b}$ (расчет по рекуррентным формулам, β =0,5)

Как показывает анализ приведенных графиков, применение рекуррентных формул несколько расширяет границы доверительного интервала, что приводит к незначительному ухудшению точности получаемых оценок наклона аппроксимирующей тренд прямой. При этом запаздывание в моменте обнаружения возрастания или убывания тренда по сравнению с методом «скользящего окна» минимально (на 1-2 такта измерения). С другой стороны, замена точных оценок сглаженными увеличивает устойчивость алгоритма в целом к аномальным выбросам измерений.

Выводы. Таким образом, использование оценок тренда, получаемых с помощью предложенных рекуррентных выражений, позволяет осуществить прогнозирование времени исполнения отдельных программ в реальном масштабе времени при минимальном потреблении вычислительных ресурсов. Кроме анализа временных характеристик, рассмотренный метод применим и для контроля других параметров вычислительного процесса, например размера свободной оперативной памяти. На основе выполняемого прогноза становится возможным адаптивное управление ходом вычислений, предусматривающее динамический переход на другие версии программных модулей. Данные версии, реализующие упрощенные алгоритмы обработки информации, могут быть реализованы исходя из требований минимального потребления вычислительных ресурсов различного типа. Перспективным для программных систем реального времени представляется использование итерационных алгоритмов, в которых, при наличии резерва времени, происходит постепенное уточнение выходной информации.

В плане дальнейших исследований предполагается проанализировать эффективность применения рассмотренного метода при различных значениях N – числа учитываемых измерений и β – сглаживающего коэффициента.

Список литературы: **1.** *Советов Б.Я., Яковлев С.А.* Моделирование систем. – М.: Высшая школа, 1998. – 320 с. **2.** *Игнатущенко В.В., Подшивалова И.Ю.* Динамическое управление параллельными вычислительными процессами на основе статического прогнозирования их выполнения // Автоматика и телемеханика. – 1997, № 5. – С. 160 – 173. **3.** *Игнатущенко В.В., Подшивалова И.Ю.* Динамическое управление надежным выполнением параллельных вычислительных процессов для систем реального времени // Автоматика и телемеханика. – 1999. – № 6. – С. 142 – 157. **4.** *Байцер Б.* Микроанализ производительности вычислительных систем: Пер. с англ. – М.: Радио и связь, 1983. – 360 с. **5.** *Липаев В.В.* Проектирование программных средств. – М.: Высшая школа, 1990. – 303 с. **6.** *Никифоров И.В.* Последовательное обнаружение изменения свойств временных рядов. – М.: Наука, 1983. – 200 с. **7.** *Назаров С.В., Барсуков А.Г.* Измерительные средства и оптимизация вычислительных систем. – М.: Радио и связь, 1990. – 248 с. **8.** *Тривоженко Б.Е.* Выделение трендов временных рядов и потоков событий. – Томск: Изд-во Том. ун-та, 1989. – 286 с.

Поступила в редакцию 15.04.04